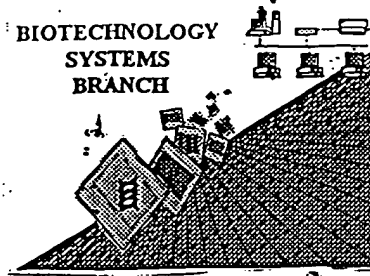


**RAW SEQUENCE LISTING**  
**ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/785,689

Source: OIPÉ

Date Processed by STIC: 11/30/01

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

**Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

## Raw Sequence Listing Error Summary

### ERROR DETECTED

### SUGGESTED CORRECTION

SERIAL NUMBER: 09/785,689

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY ETO SOFTWARE

- 1 ☒ **Wrapped Nucleics  
Wrapped Aminos** The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 ☐ **Invalid Line Length** The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ☒ **Misaligned Amino  
Numbering** The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 ☐ **Non-ASCII** The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ☐ **Variable Length** Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 ☐ **PatentIn 2.0  
"bug"** A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 ☐ **Skipped Sequences  
(OLD RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 ☐ **Skipped Sequences  
(NEW RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9 ☐ **Use of n's or Xaa's  
(NEW RULES)** Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ☐ **Invalid <213>  
Response** Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ☐ **Use of <220>** Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 ☐ **PatentIn 2.0  
"bug"** Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

Re-run

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/785,689

DATE: 11/30/2001  
TIME: 12:17:09

Input Set : A:\Seq.lst  
Output Set: N:\CRF3\11212001\I785689.raw

Does Not Comply  
Corrected Diskette Needed  
Errors on pp. 2-4

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Lindler, Luther E  
7 Warren, Richard  
8 VanDeBerg, Lillian  
9 Rubin, Fran

11 (ii) TITLE OF INVENTION: Protein From Brucella Species

13 (iii) NUMBER OF SEQUENCES: 2

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Hendricks and Associates

17 (B) STREET: P.O. Box 2509

18 (C) CITY: Fairfax

19 (D) STATE: Virginia

20 (E) COUNTRY: United States

21 (F) ZIP: 22031

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/785,689

C--> 31 (B) FILING DATE: 20-Feb-2001

32 (C) CLASSIFICATION:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Hendricks, Glenna

36 (B) REGISTRATION NUMBER: 32,535

37 (C) REFERENCE/DOCKET NUMBER: lindler

39 (ix) TELECOMMUNICATION INFORMATION:

40 (A) TELEPHONE: (703) 425-8405

41 (B) TELEFAX: (703) 425-8406

ERRORED SEQUENCES

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 1025 base pairs

48 (B) TYPE: nucleic acid

49 (C) STRANDEDNESS: single

50 (D) TOPOLOGY: unknown

52 (ii) MOLECULE TYPE: DNA (genomic)

54 (iii) HYPOTHETICAL: NO

56 (iv) ANTI-SENSE: NO

58 (vi) ORIGINAL SOURCE:

59 (A) ORGANISM: Brucella

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/785,689

DATE: 11/30/2001  
TIME: 12:17:09

Input Set : A:\Seq.lst  
Output Set: N:\CRF3\11212001\I785689.raw

```

E--> 64 CCCCTGACAT AACCCGCTTT GTCCAAATTT TTTCAACTTT TCCTGTAGGA GATTTTATGA
W--> 65      60
E--> 67 ACACTCGTGC TAGCAATTTT CTCGCAGCCT CATTTTCCAC AATCATGCTC GTCGGCGCTT
W--> 68      120
E--> 70 TCAGCCTGCC CGCTTTCGCA CAGGAGAATC AGATGACGAC GCAGCCCGCG CGCATCGCCG
W--> 71      180
E--> 73 TCACCGGGGA AGGCATGATG ACGGCCTCGC CCGATATGGC CATTCTCAAT CTCTCGGTGC
W--> 74      240
E--> 76 TACGCCAGGC AAAGACCGCG CGCGAAGCCA TGACCGCGAA TAATGAAGCC ATGACAAAAG
W--> 77      300
E--> 79 TGCTCGATGC CATGAAGAAG GCCGGCATCG AAGATCGCGA TCTCCAGACA GCGGGCATCA
W--> 80      360
E--> 82 ATATCCAGCC GATTTATGTC TATCCTGACG ACAAGAACAA CCTGAAAGAG CCTACCATCA
W--> 83      420
E--> 85 CCGGCTATTC TGTATCCACC AGTCTCACGG TTCGCGTGCG CGAACTGGCC AATGTTGGAA
W--> 86      480
E--> 88 AAATTTTGGG TGAATCCGTC ACGCTCGGTG TTAATCAGGG CGGTGATTG AACCTGGTCA
W--> 89      540
E--> 91 ATGATAATCC CTCCGCCGTG ATCAACGAGG CGCGCAAGCG CGCAGTGGCC AATGCCATTG
W--> 92      600
E--> 94 CCAAGGCGAA GACGCTTGCC GACGCTGCAG GCGTGGGGCT TGGCCGTGTG GTGGAAATCA
W--> 95      660
E--> 97 GTGAAGTGA CCGCCCGCCC ATGCCGATGC CAATTGCGCG CGGACAGTTC AGAACCATGC
W--> 98      720
E--> 100 TAGCAGCCGC ACCGGACAAT TCCGTGCCGA TTGCCGAGG CGAAAACAGC TATAACGTAT
W--> 101      780
E--> 103 CGGTCAATGT CGTTTTTGAA ATCAAGTAAA TAGCTGGGGT ATGACGCCCT TTGCCACCTG
W--> 104      840
E--> 106 ATACAAAACG CCGGCCTGGT TTCACAGGCC GGTTTTTTTG ATTAGAGCGC GTTTCGATCT
W--> 107      900
E--> 109 GATTGAATCC GATCGGCGCT CTAATCCTTT GTTTTGACGC GCATCTTTTC CGAAAACCGT
W--> 110      960
E--> 112 TTCACACTTT TCGGGATGCG GTCTAGCGGA TGATCGGGCA ACCGCGCGTA TCGGCAAATG
W--> 113     1020
E--> 115 TCACG
W--> 116     1025

```

Wrapped nucleic  
- See error summary  
sheet, item 1

```

118 (2) INFORMATION FOR SEQ ID NO: 2:
120 (i) SEQUENCE CHARACTERISTICS:
121 (A) LENGTH: 250 amino acids
122 (B) TYPE: amino acid
123 (C) STRANDEDNESS: single
124 (D) TOPOLOGY: unknown
126 (ii) MOLECULE TYPE: protein
128 (iii) HYPOTHETICAL: NO
130 (iv) ANTI-SENSE: NO
132 (v) FRAGMENT TYPE: internal
134 (vi) ORIGINAL SOURCE:
135 (A) ORGANISM: Brucella
136 (B) STRAIN: unknown

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/785,689

DATE: 11/30/2001  
TIME: 12:17:09

Input Set : A:\Seq.lst  
Output Set: N:\CRF3\11212001\I785689.raw

```

139 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
141 Met Asn Thr Arg Ala Ser Asn Phe Leu Ala Ala Ser Phe Ser Thr
E--> 142 Ile
E--> 143 1 5 10 15
146 Met Leu Val Gly Ala Phe Ser Leu Pro Ala Phe Ala Gln Glu Asn
E--> 147 Gln
E--> 148 20 25 30
151 Met Thr Thr Gln Pro Ala Arg Ile Ala Val Thr Gly Glu Gly Met
E--> 152 Met
E--> 153 35 40 45
156 Thr Ala Ser Pro Asp Met Ala Ile Leu Asn Leu Ser Val Leu Arg
E--> 157 Gln
E--> 158 50 55 60
161 Ala Lys Thr Ala Arg Glu Ala Met Thr Ala Asn Asn Glu Ala Met
E--> 162 Thr
E--> 163 65 70 75 80
166 Lys Val Leu Asp Ala Met Lys Lys Ala Gly Ile Glu Asp Arg Asp
E--> 167 Leu
E--> 168 85 90 95
171 Gln Thr Gly Gly Ile Asn Ile Gln Pro Ile Tyr Val Tyr Pro Asp
E--> 172 Asp
E--> 173 100 105 110
176 Lys Asn Asn Leu Lys Glu Pro Thr Ile Thr Gly Tyr Ser Val Ser
E--> 177 Thr
E--> 178 115 120 125
181 Ser Leu Thr Val Arg Val Arg Glu Leu Ala Asn Val Gly Lys Ile
E--> 182 Leu
E--> 183 130 135 140
186 Asp Glu Ser Val Thr Leu Gly Val Asn Gln Gly Gly Asp Leu Asn
E--> 187 Leu
E--> 188 145 150 155
E--> 189 160
191 Val Asn Asp Asn Pro Ser Ala Val Ile Asn Glu Ala Arg Lys Arg
E--> 192 Ala
E--> 193 165 170 175
196 Val Ala Asn Ala Ile Ala Lys Ala Lys Thr Leu Ala Asp Ala Ala
E--> 197 Gly
E--> 198 180 185 190
201 Val Gly Leu Gly Arg Val Val Glu Ile Ser Glu Leu Ser Arg Pro
E--> 202 Pro
E--> 203 195 200 205
206 Met Pro Met Pro Ile Ala Arg Gly Gln Phe Arg Thr Met Leu Ala
E--> 207 Ala
E--> 208 210 215 220
211 Ala Pro Asp Asn Ser Val Pro Ile Ala Ala Gly Glu Asn Ser Tyr
E--> 212 Asn
E--> 213 225 230 235
E--> 214 240
216 Val Ser Val Asn Val Val Phe Glu Ile Lys

```

Same error,  
plus misalign  
amino numbering

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/785,689

DATE: 11/30/2001  
TIME: 12:17:09

Input Set : A:\Seq.lst  
Output Set: N:\CRF3\11212001\I785689.raw

*misaligned*

E-- 217

245

250

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/785,689

DATE: 11/30/2001

TIME: 12:17:10

Input Set : A:\Seq.lst

Output Set: N:\CRF3\11212001\I785689.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:64 M:254 E: No. of Bases conflict, Input:0 Counted:60 SEQ:1  
L:65 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
M:254 Repeated in SeqNo=1  
L:68 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:71 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:74 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:77 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:80 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:83 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:86 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:89 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:92 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:95 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:98 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:101 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:104 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:107 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:110 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:113 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:116 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:142 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:0  
M:332 Repeated in SeqNo=2